# Putting the Physical into the Digital:
# Issues in Designing Hybrid Interactive Surfaces

David Kirk*, Abigail Sellen, Stuart Taylor, Nicolas Villar, Shahram Izadi

Microsoft Research Cambridge
Cambridge, UK,
CB3 0FB

{asellen; stuart; nvillar; shahrami}@microsoft.com

## ABSTRACT

Hybrid surfaces are interactive systems combining techniques of direct-manipulation multi-touch surface interaction with elements of tangible user interfaces (TUIs). The design space for such complex hands-on computing experiences is sufficiently broad that it can be difficult to make allocation of function decisions and decide when interface elements should be given either a physical or digital instantiation and the extent to which different interface functions should be made to model real-world interactions. In this paper we present two case studies of hybrid surface systems we are developing and discuss how we have reasoned through these kinds of design decisions. From this we derive a set of observations about properties of physical and digital elements, and offer them as a design resource.

## Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Design, Human Factors.

## Keywords

Hybrid surfaces, Tangibility, Direct-touch, TUI, Design.

## 1. INTRODUCTION

Over recent years there has been growing interest in tangible user interfaces (TUIs) [11, 16]. TUIs are generally taken to refer to systems where everyday physical objects are used to control and sometimes display digital information. There has been much discussion of how we might classify, categorize and otherwise define such systems [6, 3, 26], and from those that build them, there has been a tendency to search for explanations and reasons as to why TUIs might offer benefits over more conventional graphical user interfaces (GUIs). These arguments have focused on such issues as the natural affordances of physical objects [5], the embodiment available through tangibility [7], the spatial multiplexing and bimanualism [4] of tangibles, and the purported ability for experiential learning through direct manipulation [20, 22].

However recent developments pose significant challenges for current conceptions of TUIs. Chief among these is the advent of multi-touch enabled interactive surfaces or what is often called "hands-on computing". This class of device, often seen in a tabletop format (but which can be achieved through a variety of means and form factors), allows users to directly touch representations of their digital data using fingers and hands [28]. In support of this, digital objects on a surface can also be rendered to look and behave like physical objects, having 3-dimensions and operating under simulated laws of physics [1], with applicable multi-touch gestures typically modeled on the way we would physically manipulate such objects [28].
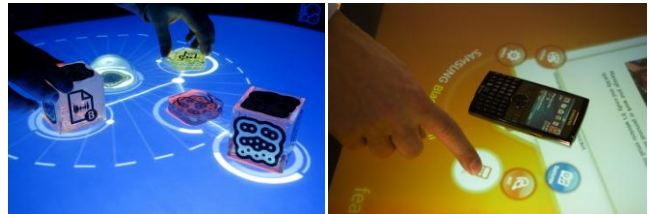


**Figure 1. Hybrid surfaces such as reacTable and Microsoft Surface allow users to directly manipulate digital objects using multi-touch input, but physical objects are also recognized and have effects in the digital world.**

These developments weaken many of the arguments for TUIs that have been put forward. For example, in Fitzmaurice et al's [5] paper, the advantages of TUIs such as encouraging two-handed interaction, allowing parallel input, making interface elements directly accessible, and affording collaborative use can all be said to be features of today's hands-on, multi-touch, computing systems. Even recent papers [12] frame the benefits of TUIs in terms of the direct control of digital representations with the hands, and spatial and conceptual coherence between input and output; benefits which are as true of multi-touch surfaces as they are of traditional TUIs. In addition to this there is no reason why the kinds of metaphor argued for by Fishkin [3], as would be presented in a tangible *physical* element, can't also be represented in a tangible *graphical* element. Also, whilst Hornecker [8] describes the embodied benefits of TUIs, direct-touch interfaces, where we have a hands-on relationship with the data, could also therefore be cast as heavily embodied interactions. So it would seem then that recent developments undermine at least some of the rationale for the benefits and power of tangibility.

Perhaps due to the relative recency of robust multi-touch technologies and their relative unavailability, amongst many interface designers there has traditionally been, and therefore remains, a strong desire to include aspects of physicality in their interfaces (often including the integration of external physical devices with the interactive surface such as mobile phones). Accordingly, many multi-touch interactive surfaces incorporate as a major aspect of their design the use of tangible elements, URP [27], Microsoft Surface and Reactable [16] (Figure 1) all being examples. These systems therefore are not simply TUIs, but what we might more appropriately refer to as "hybrid" surfaces. Reactable [16] is a good example of this: the surface on which tangible elements are manipulated is also an interactive touch surface for mixing, blending and playing with music [ibid]. Of course any interactive system such as a PC is in a sense hybrid in that it consists of both physical and digital elements. But by hybrid surface here, we mean systems where the tangible element is tightly coupled and directly mapped to an interactive (touch-sensitive) display surface, such that its activities only have an impact within the bounded, sensible region, of that interactive surface and on which direct-touch with hands is an equally viable method of interaction to tangible object-tracking.

As a research group we have been actively engaged in building hybrid surfaces. In so doing, it has become clear that there are many important design decisions that need to be made. On the basis of our own experience, the two key questions we have struggled with are:

- When should an interface designer choose physical over digital objects as tools for system functionality?
- And, for systems with direct touch capability, to what extent and in which aspects should interaction with graphical digital elements be made to emulate the physical world?

Put very simply, in designing hybrid surfaces decisions must be made about allocation of function. Here we have found relatively little guidance in the research literature which helps with making these kinds of decisions especially when building systems where both direct-touch and tangible object manipulation are available.

Herein, we present by way of illustration, two examples of projects we are currently developing: "VPlay" [25] and "Family Archive" [14]. In each case, we describe our reasoning about how and why we chose either a physical or a digital instantiation of interface elements. In the case of Family Archive, we also discuss our deliberations about the extent to which we needed to make interface interactions appear like 'real world interactions' [15]. From these examples we distill a set of observations about the more general issues one needs to consider when making these choices. Many of these come down to considering the natural affordances of physical and digital objects but also include consideration of the nature of the sensed relationship between user and machine. These observations are not exhaustive, but are driven by our own experiences and reflections during design and deployment of such systems.

We hope that these examples and the conclusions we draw from them can be used as a design resource in two ways. First, they may help to evaluate hybrid systems that already exist, and help us understand why some interface interactions work well or badly. Second, we hope this might offer design guidance for those developing such systems and reasoning through similar kinds of design decisions.

## 2. FRAMEWORKS FOR TANGIBILITY

We begin this process by first highlighting existing frameworks for tangibility which helps to place our work in context.

Notions of space- and time-multiplexed interaction were first raised by Fitzmaurice et al through their work in *graspable* user interfaces [5]. Time-multiplexed interaction is characteristic of the way a user operates a graphical interface with a mouse: using a single generic device, the user sequentially selects and manipulates virtual elements on the screen. The role of the mouse is constantly redefined over time, its function determined by the graphical context at a particular moment. In empirical studies, the advantages of time-multiplexed over space-multiplexed control were demonstrated [4]. Users performed better when operating interfaces that used dedicated controls to manipulate associated on-screen graphical objects, compared to conditions where time-multiplexed controls were used. The results confirmed benefits which had been vocalized in earlier work which proposed that "distinct controls for specific functions provide the potential to improve the directness of the user's access, such as through decreased homing time and exploiting motor memory." [2].

The results of the above body of work are often cited in the literature as evidence for the adoption of physical user interface elements. However, as we have said, recent advances in multi-touch technologies have made us question the validity of this assumption: if it is possible to realistically render specialized graphical tools on the screen which can be operated by responsive direct touch manipulation and in a space-multiplexed manner, then what are the advantages of using tangible elements?

Ullmer and Ishii proposed the "model-control-representation (physical and digital)" interaction model (MCRpd) to draw attention to the particular properties of user interfaces that make use of *tangible* interface elements [26]. In this work, the broad definition of a tangible is an artifact that clearly acts both as a *physical representative* of a digital user interface concept, and as a *tool* with which to manipulate that concept. Within this framework, a physical instantiation of a tangible can be complemented by a *digital representation*: dynamic graphics or audio closely coupled with the physical object. The vision advocated by Ullmer and Ishii is one where "atoms and bits" are closely intertwined, and control (both input and output) of intangible and transient digital information takes places through dedicated and permanent tangible artifacts [11]. This compelling vision has inspired research efforts into making interfaces more tangible.

A number of practitioners in the wider research community (including ourselves) have struggled to situate their work within a strict definition of tangibility. The relationship between digital representations and physical objects is often ambiguous. For example, some may argue that a digitizer pen is a tangible user interface element as its physical shape is closely coupled with the ability to generate a digital ink trail on a screen. Others may counter that, in fact, the pen is simply a two-dimensional input device, and is no different to a mouse in this respect. This discussion has given rise to a number of complementary frameworks, which have been proposed as means to unpack issues surrounding the role of tangibility in user interfaces and analyze the success of design which bring the digital and physical together.

In order to provide a generalized view across the design space, Holmquist et al. deconstruct tangible user interface elements into

three categories [6]. *Containers* are generic objects that can be used to move information between devices; *tokens* are a specialized form of container, which physically represents the information they are associated with, and *tools* are objects that are used to manipulate the information. Fishkin [3], contributes a taxonomy of TUIs proposing that tangibility should not be considered a binary quality of a UI that is either present or absent from a design. Instead, tangibility is conceptualized as a 2D space in which any particular design can be located. The two axes of this space are *metaphor* and *embodiment*: an interface design becomes more tangible with a stronger user perception that the state of the system is contained within the physical artifact that they are manipulating; the interface is also considered to be more tangible if the system effect of a user action is analogous to the real-world effect of a comparable action. Similarly, in Koleva's framework [19], user interfaces are categorized according to the *degree of coherence* between digital and physical objects. Hornecker and Buur structure their framework around the concept of *tangible interaction* [8] which takes into account social interaction and physical space in order to provide a broader picture of the context in which tangible user interfaces can be applied. More recently, Hurtienne et al [9, 10] have tried to answer the specific question of why tangible user interfaces are perceived as being *intuitive* to use, and to this end propose a taxonomy based on the concept of image schemas – abstract representations of recurring patterns of interaction. Jacob et al [15] unites both approaches, arguing for the analysis of tangibles in relation to the extent to which they model action in the real world combining both intuitive affordances and embodiments.

These powerful theoretical tools provide valuable insights into the role of tangibility in user interfaces, serve well to analyze related work, and help us to think about our work in context. However, in the course of our own design efforts we could only draw limited practical advice from the literature. When designing hybrid systems we regularly face design choices where interface concepts can be instantiated as physical objects, iconic digital representations, gestures, or even as realistic graphical objects that follow simulated physical laws and which can be operated via direct, multi-point interaction.

Many of the interface innovations we demonstrate in this paper are not ours alone. Some are used in other systems, such as the ability to represent information with data tiles in VPlay (see also [23]) or the use of physics in an interface in Family Archive (see also [1]). Our point is not to demonstrate novelty in design, but rather to highlight how these increasingly common strategies within design must be rationalized, especially given the many possible alternative approaches one could adopt in hybrid systems. Consequently, this paper discusses our own experiences and deliberations, and in doing so tries to move toward a more practical application of the insights derived.

## 3. CASE STUDY 1: VPLAY

VPlay is an interactive tabletop system designed to support the practice of "VJing" for both seasoned VJs and novices alike. VJing is a form of performance art that typically involves the live mixing of different video sources, the result of which is projected within a performance space such as a night club to create an engaging audio-visual experience for the audience.
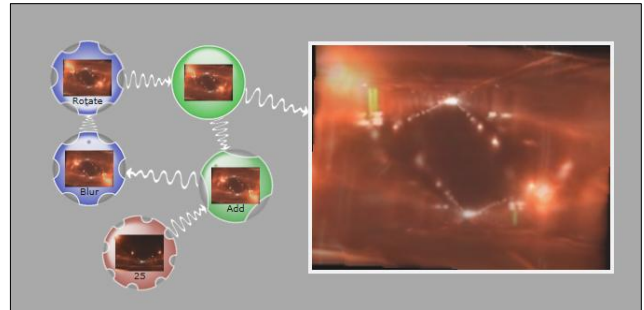


**Figure 2. The VPlay user interface. In this example, a video clip (bottom object - red) is connected to a display object (rectangular window) via various mixers and splitters (green) and effects (blue).**

Traditional VJ practice involves the use of laptops (running dedicated VJ software), video mixers and other peripheral devices. This approach allows an expert VJ to manipulate video footage in real time to produce visually stimulating outputs. However, this approach presents a 'closed' system in that it offers little opportunity for collaboration, either with other VJs, or members of the audience.

One goal in designing VPlay, therefore, was to see whether an interface on an interactive surface would open up collaborative opportunities not possible with traditional set-ups. This also involved designing the VPlay software using a simple "object-based" model, where objects on the surface could have different effects by being brought into close proximity to each other. We surmised that this might minimize learning and hence allow and encourage new users to start interacting with the system almost immediately. The design of this interface was inspired by ReacTable [16] which achieved something similar in encouraging collaborative performance with music.

The objects include video clips, mixers, splitters, effects and display windows. A simple menu system enables the creation of these objects. They can then be dragged around the interface and connected using an underlying set of rules. For example, a "video clip" object produces a single output, and has no input, an "effect" object takes a single input and generates a single output and a "mixer" object takes two inputs and produces a single output. Visual feedback appears on the surface to reinforce the nature of these connections when objects are brought near to each other. Using a combination of these objects, users are able to create new effects on one or more video streams to produce a visually appealing output. Figure 2 shows an example of the VPlay interface.

During early design phases of VPlay, many questions arose relating to interface form factor and whether to incorporate the use of tangible objects versus direct manipulation of the underlying digital objects. Practical consideration was also given to the context of use for the system and the range of potential users. These issues sometimes led to conflicting interface requirements, but also led us into deeper discussions of the nature of physical versus digital elements of the interface. As a result, the system is still evolving and its design has already undergone several iterations, such as the all-digital version shown in Figure 2.

### 3.1 Deciding on Digital or Physical Elements

Similar to ReacTable, a spatial proximity model was seen as an interesting way to determine the strength of any given effect on

the output window, where moveable icons representing effects and video clips (whose relative spatial proximity also determined the resultant mix) should be used. One important set of design decisions for VPlay was how to represent the digital "stuff" that constituted the active components of the interface, how to control them, and how to make them interact with the main viewing window containing the video output.

The choices were in fact more complex than they might at first seem. We reasoned that the objects in the interface serve essentially two different purposes: First, they act as a *means of controlling information*. For VPlay, this was true in the sense that some objects were designed to be tools for control (such as mixers and effects objects), which meant that users first needed to be able to identify these, distinguish them from one another, and sense their proximity to one another. This was also true in that tools themselves needed to be manipulated, allowing users to move them, turn them, scale them etc. A second role of the objects was as *representations of information*. Here, the way in which objects interacted with each other reinforced different mental models for users. Thus, rendering the objects as tangible or presenting them as digital, graphical objects impacts both of these functions.

With Reactable [16], a decision was made to provide physical objects for users to control and represent sounds and effects. With VPlay, it was not clear that this would necessarily be the best. In fact, as with any hybrid system, it is not simply the physical/digital choice that confronts the designer, but the coupling of the physical with the digital. In this case, we considered that the best way to experiment with tangibles was by constructing them out of transparent acrylic pieces which could be sensed by the surface, the digital content being projected through each object. This had the advantage that the corresponding content could be dynamic yet at the same time be linked to the position of each tangible object. Figure 3 shows a completely digital "effect" object, and the equivalent overlaid with a transparent physical object.
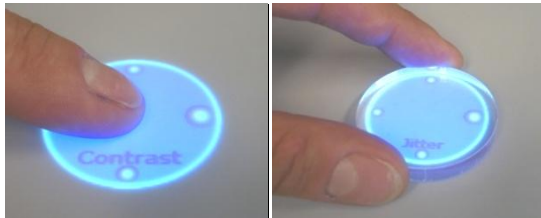


**Figure 3. An "effect" object rendered as a digital object on the surface (left) or overlaid with a piece of acrylic (right).**

Note that this design choice downplays the role of what have been called "phicons" in TUIs [11]. Phicons have been defined as physical objects that by their very nature convey their semantics, such as the use of a physical eraser to signify its ability to erase digital content, or the use of and inkpot signifying a tool to change pen colour.

### 3.1.1 Disambiguating and Identifying Objects
When discussing the pros and cons of different options we reflected on the differences in the nature of the perceptual feedback from a physical object versus a digital icon. We reasoned that a physical token offers up a diversity of tactile properties which can make it distinguishable both from its background, and also from other physical objects. The latter could be maximized by designing in variation in physical shape or texture.

Conversely, the tactile feedback from digital icons is limited, objects feeling indistinguishable from the surface. Such objects, their type, their number and their placement can therefore only be disambiguated visually in the absence of any other kind of perceptual feedback. Of course this isn't to say that with a digital icon disambiguation through tactile means is impossible: One can design in a distinct vibro-tactile response, or different auditory cues associated with different objects. But to achieve this requires significant modification and perhaps even non-trivial innovation to the existing technology. With physical tokens, in sense, all of this comes "for free".

These issues were of practical importance when we considered that VJing typically involves a combination of interacting with the VJ software user interface and watching the generated output on the projected display. At times, a VJ will focus their attention on the interface, but at other times it will either be on the projected output or on the audience. When looking elsewhere, they will often continue to manipulate the user interface to modify aspects of the mixing process and will hence be performing actions on the interface eyes-free [25].

### 3.1.2 Objects as Tools for Control
With this in mind, it is clear then that physicality may offer advantages in identifying objects on a surface, distinguishing one from another, and sensing their relative placement through a variety of senses, not just visual. But advantages might also arise in terms of greater or finer-grained, eyes-free control of objects. One example of this is the use of an object for 'scratching' (also referred to as scrubbing) video clips. Here we reasoned that a specially designed dedicated physical object provided advantages such as enabling the user to locate and maintain contact with the object on the basis solely of its tactile qualities.

Beyond the issue of eyes-free control, there might also be more fundamental issues at stake when considering objects as tools for control. Digital objects can of course be manipulated directly using multiple fingers and two hands. But physical objects require contact in a different way and in ways we naturally understand. There is a tight coupling between an action and a physical object's movement such that the result of an action is both immediate and reliably consistent[1]. We have no such assurances in the digital world: gestures and their impact on objects are defined by the designer and the developer of the systems. Effects on objects can indeed be achieved without contact at all (as new systems such as SecondLight attest [13] which act beyond the surface and allow the system to project information onto and into objects resting on or being held above the interactive surface). In some systems this might be desirable and offer interesting affordances, especially in a performance setting. For the interface we had in mind, however, the benefits of such functionality were not clear.

### 3.1.3 Objects as Representations
If physicality confers many potential benefits in terms of identifying and controlling objects on a surface, there are other affordances of the physical which we then considered would undermine their utility in our design. Digital objects can appear

---

[1] Of course a physical object on an interactive surface must still be tracked so the digital effects might lag but a user can be immediately sensible of the fact that the physical object has been performed correctly by the very act of its manipulation.

and disappear (be generated or destroyed) in ways in which physical objects simply cannot. Digital objects can occur in multiple locations simultaneously, while physical objects are inherently unique. And digital icons can be instantaneously copied in ways that physical ones cannot. Clearly then this suggested a variety of ways in which the digital offered benefits over the physical for VPlay. Part of the creative process would be supported by the ability to reproduce clips and effects, and for these objects to be generated at will.

The barriers put in place by physicality do not necessarily help here. On the one hand, the use of transparent material for physical objects gives the illusion of containment, such as when a video clip is projected up through the object. But of course this is simply an illusion. This could be further reinforced by having one-to-one mapping between a physical object and the media or effect it is linked to. For example, placing a tangible object on the surface makes the digital content appear underneath it; picking it up off the surface, makes it disappear. However, in this case, the number of clips and effects available would be limited to the number of physical objects that exist. On the other hand, if the physical object merely becomes a way of creating or displaying many different digital objects, then each physical object effectively becomes a tool or a handle for digital information rather than a container.

There are some interesting possibilities that come from the use of physical objects as containers which we also considered, however. For example, we envisioned that a physical object might "contain" a set of effects or a particularly creative piece of video that could be given to others or carried to use in other venues with similar systems. But the downside of this is also obvious: physical objects can be lost, need to be transported, will ultimately create clutter (which is anathema to an interactive surface) and so on.

Finally, we considered the nature of the dynamics and change characteristics of the physical and digital. Physical objects on an interactive surface are generally static: unmoving unless moved by the user and not easily changing shape or state[2]. Conversely, digital objects can be dynamically moved and reshaped. This led to a number of interesting possibilities, including a macro feature that enabled users to record and playback (in real time) the positions and movements of digital objects within the interface. Such recall of a layout, poses problems for a purely tangible interface, for example, once a layout is recalled digitally, the user must re-associate the physical objects with the underlying digital objects.

### 3.1.4 Summary
Taking into account all of these issues, we decided the optimal design was to render any dynamic data, such as the video and effects data, as graphical, digital objects. This would mean such objects and their interactions could be easily generated, copied,

---

[2] Here one might consider the actuated workbench [21] which allowed physical objects to move on a surface but controlled by the underlying machine (obviously such an approach requires significant technical infrastructure and would not be usable with a rear-projected infra-red optical multi-touch system such as the platform we were developing on). However, the objects themselves in the actuated work-bench, remain static in terms of shape and state – whereas with the graphical rendering of the direct-touch GUI interfaces elements can break both of these conventions.

deleted, recorded and replayed. We also decided that the role of tangibles would be limited to providing physical tools for control. These could be overlaid on top of digital objects which would continue to exist even in the absence of these tools. In other words, we have yet to find a good role for physical objects acting as containers. One of the implications of this too was to work on new kinds of physical objects so that they could be more easily distinguished on the basis of touch alone. This is the instantiation of the system we are currently working on and early field testing with digital only versions is highlighting the need for tangible control [25].

## 4. CASE STUDY 2: FAMILY ARCHIVE
In our second case study we examine similar issues of choosing between physical and digital elements, but also raise a number of new ones in terms of the nature of interaction in the digital world and the extent to which it mimics the physical world.



**Figure 4. Version 1 of the Family Archive.**

Family Archive is an interactive tabletop which allows users to capture, manage, manipulate, display and store memorabilia. It differs from many archiving systems in that it is designed to integrate digital media (e.g. photos) with digital images of physical artifacts. A motivation for this was the archiving of legacy collections of print photos but also the potential value of capturing digital traces of other physical objects for which families have sentimental attachment such as children's artwork, a baby's first pair of shoes, souvenirs from family holidays and so on. The resulting system consists of a multi-touch surface with an overhead camera for scanning as shown in Figure 4. We are still iteratively evolving the design of the system, as a result of a field trial with real households [14].

## 4.1 Deciding on Digital or Physical Elements
A key element of the user interface utilises the metaphor of boxes to allow users to create loose collections of different digital objects. We decided that the interface should allow users to create new boxes, put collections of objects into them, label the boxes, empty them out, and even rummage through their contents. This came from some early fieldwork where we frequently saw people use physical boxes to store heterogeneous sets of items, containing them in a relatively unstructured way. Consequently this felt like a natural metaphor that households would understand.

How we implemented the use of boxes within the archive raised many different design possibilities however. As with VPlay, we considered the pros and cons of digital versus physical representations of boxes. Most physical options were likely to require more complex implementation. Therefore we felt we needed some strong reasons to justify the use of tangibles in the interface.

### 4.1.1 Controlling and Manipulating Boxes
Unlike with VPlay, the need to disambiguate and manipulate boxes in an eyes-free manner seemed unlikely. Fieldwork looking

at people's interactions with photos [18] and other sentimental artifacts [17] has shown that people very much focus on the materials before them when they organize them, rummage through them, and share them. Therefore this aspect provided no strong argument for the use of physical boxes.

When it came to the issue of control, however, the picture was quite different. We wanted users to be able to move boxes around the virtual environment, open them up to view their contents, tip them over to spill their contents, put objects into boxes, and close them to keep the contents safe. All of these actions in the physical world are complex, but are nonetheless intuitive to us. We have a lifetime of skills in manipulating physical containers in this way. We could therefore exploit this fact, and use physical models of boxes to emulate this, linking the actions on these physical objects to the behavior of content on the surface. So for example, placing boxes on the surface and tipping them, could "spill out" associated collections of media. Physical boxes could also be opened up or closed, having implications for whether digital contents could be spilled, viewed and so on. Using tangibles in this way would essentially amount to using them as tools for rich intuitive control of virtual boxes, requiring some way for the user to specify which virtual box is to be acted upon.

### 4.1.2 Representing Boxes
An additional possibility is that these same physical boxes are designed such that they act as the "containers" for the digital media they control. In other words, the physical objects could be exclusively linked to collections of digital media, similar to the example in VPlay where linked digital content appears on the surface only when a specific object is placed on it, and disappears when the object is removed. We also considered reinforcing this idea by allowing users to view linked digital content through the boxes, perhaps by building them with a transparent bottom. The boxes would then essentially be used as physical frames for viewing associated digital content. Likewise, digital objects could be collected into boxes by moving the physical frame over them, or by flicking the objects towards it. Each physical box would therefore represent a given collection of digital data, and at the same time be a means of gathering items together.

Extensibility was a problem with this approach. Here we were concerned that there would be a limited amount of space surrounding the archive device in which real boxes could be stored. Furthermore, users would be also restricted by the number of boxes given to them, as it was difficult to decide how users themselves could make new ones. Digital boxes, on the other hand, could be generated easily and almost without limit, giving users far more flexibility.

Beyond extensibility there was also a problem stemming from what the surface might understand of an interaction. A clear problem with physical boxes was the way in which they would need to be tracked by the surface system to facilitate a variety of user led actions. Whilst the physical shape of a box clearly affords picking up and tipping actions the surface technology we were developing on (rear-projection interface with infrared optical tracking for multi-touch) is limited in its sensing capabilities, being confined to 2D interaction in contact with its surface This strongly implies that the system needs to have more tracking technology to fully support 3D interactions with objects above the surface to really maximize the benefit of having such 3D action possible. In many ways the lack of functionality, such as having 6 d.o.f. tracking available in a Vicon system or from the AR toolkit,

constrained our decision to opt for 3D boxes. A key part of making these design decisions is grounded in deciding when it makes sense to add more hardware and complexity to the system and when to find cheaper and easier software solutions which capture most of the intent and user demands at a fraction of the cost (not just financial but developmental and reliability costs).



**Figure 5. Different means of making boxes (left pair) and different means of inking on boxes (right pair).**

### 4.1.3 Creating New Boxes
A final alternative we considered was therefore to consider tangible boxes not as containers but as tools for creating new boxes. Here in fact we began to experiment with two versions of this (see Figure 5), one where the user simply touches a "new box" icon and one appears, and a different version where the user picks up a small physical model of a box marked "new box" and a digital box appears where the box makes contact with the surface.

One advantage of the physical approach is that it removes the need for a graphical icon on the touch surface, clearing more space for interaction. Another is that the action of making contact with the screen with the tangible object at once initiates the command to generate a new box and specifies the location of the box, unlike the digital option. This then seemed an interesting tool where users could use the physical box almost as a stamp to create multiple instances of digital boxes.

But there were also drawbacks: if we were to deploy this important object in real houses with children, would it sometimes go missing? Physical items can be detached, become lost or otherwise be misappropriated. Any argument in favour of this technique seemed outweighed by these concerns.

### 4.1.4 Opting for Digital Boxes
After considering all of these issues we decided that there were strong reasons to opt for entirely digital ways of manipulating and representing boxes in the interface. While physical representations might afford more natural manipulations, we felt these arguments were not as strong the need to quickly create multiple containers that could be linked to particular collections of digital objects. We were also concerned that there might be additional problems with introducing physical objects to cover only some aspects of their functionality and not others, and that this would lead to too much complexity in the design.

### 4.1.5 Inking
Interestingly, the decision we reached about boxes can be contrasted with our reasoning about whether to use a physical pen for the annotation of boxes. In our current version, users touch a pen icon to initiate "inking mode" (Figure 5). Users can then use their fingers to write on boxes as a way of labeling them. One of the problems we have noticed with the interface as it now stands is that users frequently make "mode errors" where they begin to navigate while still in inking mode, or try to ink while in navigation mode.

An option here that we began to consider was to implement instead a physical pen which, when removed from its "stand" on the physical surface alongside the touch surface, automatically puts the system into inking mode (Figure 5). Replacing it in the stand turns off inking mode. We predicted that this would create *fewer* mode errors. Research suggesting [24] that user-maintained, kinesthetic feedback such as the act of holding the pen might be effective in preventing mode errors. Of course picking up a pen doesn't mean that one is necessarily writing, in addition pens are often returned to places other than docks (such as mouths and behind ears) whilst the user task switches. So whilst tangible pen use for task switching might offer some benefits it is also again evident that there is a problem that needs to be resolved for a surface to have a more complete model of the real world, understanding the difference between a user's finger and their pen, thereby allowing the system to gracefully determine the user's intentionality.

There are other good reasons for opting for pen input for inking. In contrast to the use of boxes in the interface, the pen is quite clearly a tool with a specific function. Here we predicted that the particular affordance of a physical pen in terms of fine-grained control was much preferable to inking with one's finger. It is an artifact ideally designed with its function in mind.

In addition, unlike boxes, the pen is only a tool[3]. Boxes can be viewed as tools for manipulation of content, but they are also representations of containers. As we have seen, it is this aspect which creates problems when adopting the tangible option. With the pen, issues about multiple instances and extensibility did not arise.

Finally, we felt the valuable affordances of the physical option outweighed the possibility that the object might go missing, something we intend to further prevent by providing a clear place or dock where the pen should "live" on the Archive's physical surface. As a result, unlike the case of boxes, we plan to implement in the next version of the Archive a physical pen for annotation to replace the current digital pen "mode".

## 4.2 Modelling the Physical World in the Digital Domain

We conclude this case study by considering further implications of the decision to use purely digital representations for the notion of boxes in the interface. Physical boxes afford actions we clearly understand through a lifetime of experience interacting with them. Furthermore, these manipulations occur in a three-dimensional world. For the design of the digital domain, we needed to determine the extent to which the virtual boxes should act like real physical boxes, and the extent to which our acting upon them should or could emulate the physical world given the limits of our surface technology.

### 4.2.1 Choosing a 3D World

One important decision was whether to implement a notion of virtual 3D boxes contained within a physics-enabled virtual

world, enabling users' direct multi-touch interaction with those virtual objects. Or alternatively we could opt for simple animations of boxes that had more of an iconic feel and less of an approximation of real world boxes. For example, the interface could rely more on conventions borrowed from GUIs, such as requiring simple taps on boxes to open them up. Direct touch would still form the basis of the input vocabulary, but it would not draw on real-world manipulations or metaphors to do this.

We decided to implement a 3D physics-enabled world mainly because we believed that pushing this interactional model would help us better understand what natural features of the physical world we might be missing in any given interaction as we tried to emulate it in the digital. We further wanted to explore this in the context of a system designed to support real household archiving activities.

Accordingly we used a games engine to build a virtual world in which boxes are created, opened, filled, closed and labeled, and in which the contents of boxes could be interacted with in similar ways that *simulate* interactions between real objects. Notions of gravity, friction, inertia and so on are built in. In doing so, this determines how objects interact with each other, and how we interact with those objects (see Figure 6).The successful use of physics in an interface has been demonstrated in such work as BumpTop [1], but obviously their system differs dramatically in that it is a single-touch system designed for pen-input, consequently many of their design solutions pertain to ways of dealing with the limits of such relatively impoverished input. With our multi-touch – hands-on environment there was a much richer means for building direct-touch interactions with objects.



**Figure 6. A snapshot of the 3D world showing how boxes can be opened up and the contents viewed.**

### 4.2.2 The Problems of 3D Interaction on a 2D Surface

Clearly all of the features of perceptual feedback discussed earlier are missing from the digital emulation, in particular any tactile disambiguation between a manipulable object and its background was not possible. But this was not perceived to be problematic. A visual rendering of the objects was deemed sufficient for coherent use by the users. And the issues of generic identity and dynamics and change conferred by use of the digital boxes were all considered to be highly beneficial for rich interactions.

There was one area of natural propensity for the physical though whose absence did cause some intriguing problems. This was the loss of ability to manipulate objects in 3D space and basically the literal lack of existence of real digital boxes in the real third dimension.

By emulating 3D boxes but only really having the means to interact with them in 2D (by virtue of using a touch interactive surface as the input device) we had created an interactional

---

[3] Having said that the burgeoning numbers of pen-based interfaces such as [1] are now pushing digital pens to provide a variety of interface manipulations that are semantically incoherent with our understanding of a 'pen' such as the ability to lasso, grab and flick objects. But such an approach stands in contravention to the suggested push towards reality-based interfaces (one of the supposed advantages of tangible systems) discussed in [15].

problem. Thinking about it logically highlights the problem: How would you, for example, tip over a 3D virtual box if you can only touch it in 2D? Beyond this, any number of what might otherwise appear to be quite natural actions with a physical box such as putting things inside it or taking them out or perhaps even stacking the boxes, becomes quite a challenge when one can't grasp the box in three dimensions. Much of the intuitive understanding of the physical world and how to interact with it becomes lost when objects are emulated in a digital realm.

Ultimately, we were forced to derive solutions to these 3D interactional problems by breaking the laws of physics and creating our own set of laws or workarounds much as seen in Agarawala & Balakrishnan's [1] call for *polite physics* and the *disable physics as necessary* guideline they used. For picking a box up, we created buttons on the screen, which, when pressed, lifted or lowered the box in the z-axis (allowing us to move the box between 'floors' of our virtual archive), the other hand being used on the box itself to determine its placement on the x-y plane. To get objects into a box we designed the interaction such that, when open, boxes could "suck in" proximally adjacent digital content. This meant users could "hoover up" items by moving the box near them or over them. To get the contents out of boxes, we implemented a gestural interaction, whereby a two-fingered action at the top of the box tipped them over. This involved changing the underlying friction with the virtual surface resulting in the fact that boxes touched in this way were more prone to tip over. However, such an action requires skill to learn, and the amount of effort required to carry out such a key action was brought into question. In response, we ended up having to implement an additional shortcut by creating a button that automatically spilled box contents in a more efficient way.

### 4.2.3 Revisiting Our Decisions

Having implemented this physics-enabled world and experienced first-hand the many problems it created for manipulation of key elements of the interface, it caused us to reassess our original rationale. Originally we had judged that the benefits of manipulation conferred by tangibles were outweighed by other considerations which made us lean toward a digital solution. In fact, although we emulated some of the physical *functionality* of boxes in our digital instantiation, it wasn't necessarily enough because we hadn't had the means to digitally emulate and support all of the physical *interactions* that might normally be used with them. To exacerbate this, the physics world reinforces for users a mental model that suggests that these manipulations will be possible.

Further important issues were raised in the first field deployment of this system. As we outline in detail elsewhere [14], even when gestural manipulations did work well and were true to users' expectations, the time and effort—or almost physical "work" required to execute actions in the interface—was often seen by households as burdensome. After all, if this is a digital system, shouldn't it circumvent that work?

As we had hoped in choosing this approach, implementing the solution in this way brought to light the many ways in which emulation of a 3D world on a 2D surface presents design challenges for interaction. Going forward we need to consider new ways to optimize and integrate the affordances of the digital and the tangible in this system. We are currently considering how to, for example, build more physical controls into it to compensate

for problems of 3D manipulation. We are also building in more "digital shortcuts" to minimize user effort.

Consequently it would appear that when interactions are actually made more real-world as advocated in [15], and as can be seen in the progression of our interface over BumpTop [1] (after all having multi-touch and hands-on interaction is inherently more real-world than using a pen-based input), the increased real-world nature does not necessarily make the system more usable as suggested [see 15].

## 5. LESSONS FOR DESIGN

The case studies we have outlined, VPlay and Family Archive, make clear that there are broadly two important sets of decisions to be grappled with in designing interactive multi-touch systems when considering the extent to which we draw on the physical world:

- *Choice of Objects:* One set has to do with the choice of physical or digital elements as tools or objects that provide key features of the interface. This, as we have seen, determines the ways in which we interact with these elements, physical and digital objects having affordances which both constrain and make possible certain kinds of action. Other implications here are practical, which is especially important for systems which are designed to be deployed in the real world and not limited to demonstrations at conferences or in laboratories.

- *Emulating the Physical World:* The second has to do with the emulating the physical world within the digital domain. This encompasses how digital objects behave within the digital environment, such as whether there is a notion of physics guiding users' understanding of that world. This also, as we have seen, has implications for how we design users' interactions with those objects.

When we look to the literature on tangible interfaces, we find little guidance in making the above decisions. Much has been written about the advantages of TUIs, for example, but this analysis seems retrospective, providing little or no insight into the design process. Other systems make no attempt to provide a design rationale, but rather justify themselves implicitly in their very existence by providing us with novel and engaging experiences.

But another approach, which we advocate on the back of our experiences in building deployable systems, is to make these decisions after deep consideration of the alternatives. We have seen that one must sometimes balance the creation of a novel and engaging user experience with what makes sense for the envisioned tasks. For example, we might create a new and compelling user interface by incorporating physical objects for key features of the system, but ultimately this novelty begins to wear thin in real use. Likewise, emulating the physical world in the digital domain may lead to playful and intuitive use, but ultimately users may desire efficient shortcuts and learned conventions that reduce the need for physical work in the digital realm.

As a result of our experiences in designing the two systems we have described, we summarize the issues we confronted and the rationale we used. As should be evident, choices between physical and digital options go beyond claims of spatial multiplexing, direct touch and the support of gesture, since we take these as a

given in both the physical and digital domain in the context of today's interactive surfaces.

**Nature of the physical objects:** One question that confronts the designer is whether or not the physical objects (and indeed digital ones too) convey any semantic meaning as to their function (as in phicons). In the case of VPlay, we not only opted for generic tangibles, but transparent ones, allowing display through the objects and thus the reassignment of meaning through digital content. Whether objects need to dynamically change their semantics or create the illusion of containment are considerations here.

**Eyes-free control:** An important issue has to do with the degree to which system will require eyes-free control. If the nature of the task or environment dictates it, these are good arguments for the use of physical objects for identifying and controlling data. Otherwise, digital control elements will need to provide added non-visual feedback to make this possible.

**Affordances of physical tools:** Another issue to consider is whether there is a case to be made that physical artifacts already exist which are uniquely designed to provide the affordances for rich or accurate control (such as a pen for inking). If so, it may be worth the effort to find ways to incorporate this into the design.

**Affordances of digital tools:** On the other hand, there will be actions in the digital realm for which no obvious physical tool exists. This includes the actions of generating, reproducing, replaying, merging and deleting. For such actions, it makes sense either to rely on established digital conventions or to invent new ones.

**Controlling a 3D world on a 2D surface:** Constructing a 3D world on an interactive surface can create a design conundrum. On the one hand it encourages and suggests the use of gestures to manipulate the objects within it. At the same time, the lack of the third dimension makes many such actions impossible. Designers must be prepared to create new laws that often break the laws of physics. Indeed the design challenge here is to come up with a coherent gestural vocabulary and set of actions for users. In addition, designers also need to strike a balance between actions which are perceived as physical "work" and those that provide efficient shortcuts for users that they have grown to expect in the digital world.

**Incorporating 3D manipulation:** A way of circumventing problems caused by simulating 3D manipulation on a 2D surface is to link physical objects to the digital world. For example, tipping a physical box could result in tipping a virtual box in the corresponding digital world. The challenge here is to reconcile activities in the two disjointed worlds. For example, although they may act as one in some sense, any given box then has two identities: one physical, the other digital. An alternative is to design the system to reinforce the illusion that the worlds are connected by, for example, tipping the physical box on the surface and displaying only the contents digitally. Such an approach has some drawbacks, as is illustrated in the next four points.

**The 3D world from a 2D techno-centric viewpoint:** When creating a 3D interaction using a physical element it is important to consider how the machine is equipped to sense the interaction. Most interactive surfaces can essentially 'see' in 2D and whilst the nature of a physical object will afford to the user a variety of possible 3D manipulations the machine might not be equipped to respond adequately. So whilst physical objects can constrain action they can also open up interactions in ways which the technology cannot support, without giving it models of the real world. Provision of such models through 3D tracking or by installing 'smarts' into the tangibles, mean that the machine can do a better job of inferring the intentionality of the user as they manipulate an object.

**Physical objects as containers:** In both VPlay and Family Archive, we considered how to reinforce the illusion that physical objects were in fact "containers" of digital media. This meant, for example, designing the system such that digital media would appear when key objects were brought near or tipped onto the surface. To maintain this illusion requires that there is a one-to-one mapping between physical objects and associated digital content. For systems where many such objects are required (such as the boxes in our Archive system) this presents an impractical solution because of clutter, the need to transport the items etc. In this case, digital alternatives may be the better option.

**Loss of physical objects**: A very real issue for deployments of systems in the real world is the risk of key physical elements becoming detaching, destroyed or lost.

**Mode errors and physical feedback:** A final consideration that may sway designers toward the use of a physical object is not only that physical tools can reduce digital clutter, but that it may remove the need for modes. Interacting with a physical object can be equivalent to an implicit, user-maintained mode. The kinesthetic feedback from this can prevent mode errors.

# 6. CONCLUSIONS

This paper has presented two case studies of the design of hybrid surface technologies. In VPlay, the main issues of concern were the allocation of function in the system, and getting the balance of physical and digital elements right. In Family Archive, we confronted many of the same issues, but in also focused on the laws of digital interaction that were being developed. Principally the concern was for the extent to which the real-world should be emulated by the digital interactions.

Both experiences gave rise to a set of considerations which we have derived, as it were, from the "ground up". We have offered them here contribution to the practical development of future systems, and hope that additionally this work encourages a deeper understanding of the potential for inventing not only compelling but useful hybrid systems.

# 7. ACKNOWLEDGMENTS

# 8. REFERENCES

[1] Agarawala, A. & Balakrishnan, R. (2006) Keepin' It Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '06. ACM, 1283-1292

[2] Brown, E., Buxton, W. & Murtagh, K. (1990). Windows on tablets as a means of achieving virtual input devices. In *Proc. of INTERACT '90*, 675-681

[3] Fishkin, K. P. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*. 8, 5, 347-358.

[4] Fitzmaurice, G, & Buxton, W. (1997) An empirical evaluation of Graspable User Interfaces In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '97. ACM,, 43-50

[5] Fitzmaurice, G, Ishii, H., & Buxton, W. (1995) Bricks: Laying the Foundations for Graspable User Interfaces In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '95. ACM,, 442-449

[6] Holmquist, L., Redström, J. & Ljungstrand, P. (1999), Token-Based Access to Digital Information. In *Proc. of HUC'99*. 234-245

[7] Hornecker, E. (2005) A Design Theme for Tangible Interaction: Embodied Facilitation. In *Proc. of ECSCW'05*, 23-43

[8] Hornecker, E. & Buur, J. (2006). Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '06. ACM, 437-446

[9] Hurtienne, J. & Israel, J. H. (2007). Image Schemas and Their Metaphorical Extensions – Intuitive Patterns for Tangible Interaction. In *Proc. of the 1st international Conference on Tangible and Embedded interaction*. TEI '07. ACM, 127-134

[10] Hurtienne, J., Stössel, C. & Weber, K. (2009) Sad is Heavy and Happy is Light Population Stereotypes of Tangible Object Attributes. In *Proc. of the 3rd international Conference on Tangible and Embedded interaction*. TEI '09. ACM. 61-68

[11] Ishii, H. & Ullmer, B. (1997) Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '97. ACM,, 234—241

[12] Ishii, H. (2008). Tangible bits: beyond pixels. In *Proc. of the 2nd international Conference on Tangible and Embedded interaction*. TEI '08. ACM. xv-xxv.

[13] Izadi, S., Hodges, S., Taylor, S., Rosenfeld, D. Villar, N., Butler, A., & Westhues, J. (2008) Going Beyond the Display: A Surface Technology with an Electronically Switchable Diffuser. In *Proc. of the 21st Annual ACM Symposium on User interface Software and Technology*. UIST '08. ACM. 269-278

[14] Kirk, D. S., Izadi, S., Sellen, A., Taylor, S., Banks, R. & Hilliges, O. (2009) Opening up the Family Archive. *Paper in submission to CSCW 2010*.

[15] Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '08. ACM,, 201- 210

[16] Jordà, S., Geiger, G., Alonso, M., & Kaltenbrunner, M. (2007). The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proc. of the 1st international Conference on Tangible and Embedded interaction*. TEI '07. ACM. 139-146

[17] Kirk, D. & Sellen, A. (2008) On Human Remains: Excavating the Home Archive. *Microsoft Technical Report*. MSR-TR-2008-8

[18] Kirk, D., Sellen, A., Rother, C., & Wood, K. (2006). Understanding photowork. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '06. ACM, 761-770

[19] Koleva, B., Benford, S., Kher Hui Ng & Rodden, T. (2003) A Framework for Tangible User Interfaces. In "Physical Interaction (PI03) - Workshop on Real World User Interfaces", at *Mobile HCI '03*

[20] Manches, A., O'Malley, C. & Benford, S. (2009) Physical Manipulation: Evaluating the Potential for Tangible Designs. In *Proc. of the 3rd international Conference on Tangible and Embedded interaction*. TEI '09. ACM. 77-84

[21] Pangaro, G., Maynes-Aminzade, D., & Ishii, H. (2002) The actuated workbench: computer-controlled actuation in tabletop tangible interfaces. In *Proc. of the 15th Annual ACM Symposium on User interface Software and Technology*. UIST '02. ACM. 181-190

[22] Parkes, A., Raffle, H., & Ishii, H. (2008). Topobo in the wild: longitudinal evaluations of educators appropriating a tangible interface. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '08. ACM,, 1129-1138

[23] Rekimoto, J., Ullmer, B., & Oba, H. (2001) DataTiles: a modular platform for mixed physical and graphical interactions. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '01. ACM. 269-276

[24] Sellen, A.J., Kurtenbach, G.P., & Buxton, W. (1992). The prevention of mode errors through sensory feedback. *Human-Computer Interaction*, Vol. 7(2), 141-164

[25] Taylor, S., Izadi, S., Kirk, D., Harper, R., and Garcia-Mendoza, A. 2009. Turning the tables: an interactive surface for vjing. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems* CHI '09. ACM, 1251-1254

[26] Ullmer B. & Ishii H. (2000) Emerging frameworks for tangible user interfaces'. *IBM Systems Journal* 39(3-4), 915-931

[27] Underkoffler, J. and Ishii, H. (1999). Urp: a luminous-tangible workbench for urban planning and design. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems.* CHI '99. ACM,, 386-393

[28] Wilson, A.D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., Kirk, D. (2008) Bringing Physics to the Surface. In *Proc. of the 21st Annual ACM Symposium on User interface Software and Technology*. UIST '08. ACM. 67-76